

RISM-MOL: User Manual

Computational Chemistry and Biophysics group
Max-Planck-Institute for Mathematics in the Sciences

May 16, 2010

Contents

1	Introduction	3
1.1	How to read this manual	3
1.2	What can be calculated with RISM-MOL	3
2	Theory	3
2.1	RISM model	3
2.2	Closure relations	4
2.3	HNC, GF, PW ,KH Hydration Free Energy expressions	4
2.4	Excluded volume calculations and PWC expression	6
2.5	Bridge correctioin and HNCB expression	6
3	Algorithm	6
3.1	RISM equation reformulation	6
3.2	Short-range RISM formulation	7
3.3	Fourier iterative algorithm	8
3.3.1	Iterations with the λ coupling parameter	8
3.3.2	Fourier iteration step	8
3.3.3	Accuracy Theshold	9
3.4	Multi-Grid	10
4	Algorithm Parameters	11
4.1	Site-site potential description	11
4.1.1	Lennard Jones Mixing rules	12
4.2	Parameters of solvent	12
4.3	Parameters of the solute molecule	12
4.3.1	Format of input solute structure file	13
4.3.2	Example of the input structure file	13
4.4	UserParameters file	13
4.4.1	Water parameters	14
4.4.2	Model parameters	14
4.4.3	Numerical Scheme parameters	15
5	Program interface: startRISM function	15
5.1	Input Parameters of startRISM function	17
5.1.1	fname	17
5.1.2	input_units,output_units	17
5.1.3	Supported units	17
5.1.4	How to define input/output units	18
5.1.5	Extra parameters: xparam	18
5.1.6	Example of using extra parameters	18
5.2	Output parameters of startRISM function	19
5.2.1	Mu: Hydration Free Energy and excluded volume	19
5.2.2	Example of Hydration Free Energies calculation	19
5.2.3	Radial distribution functions: <i>g</i> and <i>grid</i> parameters	19
5.2.4	Example of calculating RDF functions	20
5.2.5	Solutions of the RISM equations: <i>gamma,c</i> and <i>potLong</i> parameters	20
5.2.6	Example of working with gamma, c , and potLong parameters	21

1 Introduction

1.1 How to read this manual

This manual is organized in the following way: In the section 2 (*Theory*) some basics of the RISM theory is given. In the subsection 2.1 some basics of RISM model are discussed, and then in the sections 2.2, 2.3 - different closure relations and expressions for calculating Hydration Free energy are described. Even if you are not familiar with RISM model, or do not have appropriate mathematical background - it would be better to have a brief look to sections 2.2, 2.3, because otherwise some options of the program will seem to you ununderstandable.

In the section 3 the RISM-MOL solver is described. The section contains some technical details of RISM calculations, and a brief description of the multi-grid technique. If you are not interested in the mathematical technicalities, you may skip this section. However, we suggest you to read subsections 3.3.1 (λ coupling) and 3.3.3 (Accuracy threshold), because in these sections are defined the algorithm parameters λ and ε , which you can change to affect the convergence rate of the algorithm and accuracy of calculations.

In the section 4 parameters of the RISM-MOL are described. In the subsection 4.1 you will find information about the potentials used in the RISM-MOL, and Mixing rules. In the subsection 4.3 you will find the information about the solute molecule parameters and useful instructions about the preparing input file for the RISM-MOL program. In the section 4.4 you will find the list of available parameters in the file *UserParameters.m*.

The section 5 describes the interface to the startRISM function (the main function in the program). There you will find the information about input/output parameters of the function and short instructions, how to start the calculations.

Note 1 *This manual contain some examples of the scripts. BE CAREFUL while copying these examples: while copying from the pdf file underscore symbols ('_') MAY LOST!. As well, BE CAREFUL about the quote symbols (''), because they can convert to some UNREADABLE FOR MATLAB quotes. CHECK your scripts, before running them in matlab.*

1.2 What can be calculated with RISM-MOL

RISM-MOL is the RISM solver of the RISM equations for the aqueous infinitely diluted solution. In the current version the temperature and the density are fixed to be $T = 300K$ and $n = 33.7$ molecules/ \AA^3 respectively. RISM-MOL allow to calculate Hydration Free energies of an arbitrary solute molecule, by six different expressions (HNC,HNCB,GF,PW,PWC,KH) (see section 2.3 for their description). Also RISM-MOL calculates the solute excluded volume. As well, with RISM-MOL you may find the site-site Radial Distribution functions between the solute atoms and the Water Oxygen and Hydrogen. RISM-MOL have an flexible settings, so you may change force-field parameters of solute and water, water model, closure relations, accuracy of calculations.

2 Theory

2.1 RISM model

In the RISM approach both the solute and the solvent molecules are treated as sets of sites with spherically-symmetric properties. Three types of site-site correlation functions are considered in the RISM: intramolecular correlation functions, total correlation functions and

direct correlation functions. For the two sites, s and s' of one molecule, the intramolecular correlation function is:

$$\omega_{ss'}(r) = \frac{\delta(r - r_{ss'})}{4\pi r_{ss'}^2} \quad (1)$$

where $r_{ss'}$ is the distance between the sites and $\delta(r - r_{ss'})$ is the Dirac delta-function. Total correlation functions $h_{s\alpha}(r)$ and direct correlation functions $c_{s\alpha}(r)$ are defined for each pair of solute and solvent sites (s and α , respectively). The total correlation functions can be expressed as $h_{s\alpha}(r) = g_{s\alpha}(r) - 1$, where $g_{s\alpha}(r)$ is the radial distribution function of solvent sites around the solute sites. Bulk solvent total correlation functions $h_{\alpha\xi}^{\text{bulk}}(r)$ are also considered, and represent the distribution of sites ξ of solvent molecules around the site α of the selected solvent molecule. Direct correlation functions $c_{s\alpha}(r)$ are calculated using the set of RISM equations for the case of infinitely diluted solution [1]:

$$h_{s\alpha}(r) = \sum_{s'\xi} \langle \omega_{ss'} * c_{s'\xi} * [\omega_{\alpha\xi}^{\text{bulk}} + \rho h_{\alpha\xi}^{\text{bulk}}] \rangle(r) \\ s = 1, \dots, N_{\text{solute}}, \alpha = 1, \dots, N_{\text{solvent}}, r \in [0; \infty) \quad (2)$$

Here $\langle x * y \rangle(r)$ is the radial part of the spherically symmetric three-dimensional convolution $\langle x * y \rangle(r) = \int_{R^3} x(\mathbf{r} - \mathbf{r}')y(\mathbf{r}')d\mathbf{r}'$, and ρ is a number density of the bulk solvent. To complete the set of RISM equations, one needs to use a closure relationship, which has the general form:

$$c_{s\alpha}(r) = e^{\Xi_{s\alpha}(r) - B_{s\alpha}(r)} - h_{s\alpha}(r) + c_{s\alpha}(r) - 1, \quad (3)$$

where $\Xi_{s\alpha}(r) = -\beta u_{s\alpha}(r) + h_{s\alpha}(r) - c_{s\alpha}(r)$, $u_{s\alpha}(r)$ is the atom-atom potential, $B_{s\alpha}(r)$ is a so-called *bridge function*, [2, 1] $\beta = 1/k_B T$, k_B is the Boltzmann constant, and T is the temperature.

2.2 Closure relations

The analytical expression for the bridge function $B_{s\alpha}(r)$ contains summation of infinite series of integrals with growing dimensionality, thus it is practically uncomputable. Different bridge approximations are used. They produce different approximations of the closure relation (3). In the RISM-MOL program two closures are implemented:

1. Hyper-Netted-Chain (HNC) closure ($B(r) \equiv 0$):

$$c_{s\alpha}(r) = e^{\Xi_{s\alpha}(r) - B_{s\alpha}(r)} - h_{s\alpha}(r) + c_{s\alpha}(r) - 1 \quad (4)$$

2. Partially-Linearized HNC closure:

$$c_{s\alpha}(r) = \begin{cases} e^{\Xi_{s\alpha}(r) - h_{s\alpha}(r) + c_{s\alpha}(r) - 1} & \Xi_{s\alpha}(r) < 0 \\ -\beta u_{s\alpha}(r) & \Xi_{s\alpha}(r) > 0 \end{cases} \quad (5)$$

Note 2 *It should be mentioned, that RISM equations with Hypernetted Chain closure did not converge for many molecules, while Partially-Linearized Hypernetted Chain closure usually show much better stability [3, 4, 1].*

2.3 HNC, GF, PW ,KH Hydration Free Energy expressions

Within the RISM theory, there are several expressions which allow one to obtain values of the hydration free energy from the total and direct correlation functions $h_{s\alpha}(r)$ and $c_{s\alpha}(r)$. In the RISM-MOL program such expressions are implemented:

Table 1: Closure - ΔG expression correspondance

Closure	Expressions
HNC	HNC
PLHNC	KH
HNC with Bridge Correction	HNCB
<i>No specific</i> (Any closure is ok)	GF, PW

- Hypper Netted Chain (HNC) [5]
- Hyper-Netted-Chain method with repulsive bridge correction (HNCB) [6]
- Gaussian Fluctuations approximation (GF)[7, 8, 9]
- Partial Wave expression (PW) [10, 8]
- Kovalenko-Hirata expression for the PLHNC closure (KH)[11, 12, 1]
- Partial Wave expression with excluded volume corrections (PWC) [13].

Below are given formulae for HNC, GF, PW and KH expressions. HNCB and PWC expressions one can find in the next subsections.

$$\Delta G_{HNC} = 2\pi\rho kT \sum_{s\alpha} \int_0^\infty [-2c_{s\alpha}(r) - h_{s\alpha}(r)(c_{s\alpha}(r) - h_{s\alpha}(r))] r^2 dr \quad (6)$$

$$\Delta G_{GF} = 2\pi\rho kT \sum_{s\alpha} \int_0^\infty (-2c_{s\alpha}(r) - c_{s\alpha}(r)h_{s\alpha}(r)) r^2 dr \quad (7)$$

$$\Delta G_{PW} = \Delta G_{GF} + 2\pi\rho kT \sum_{s\alpha} \int_0^\infty \tilde{h}_{s\alpha}(r)h_{s\alpha}(r)r^2 dr \quad (8)$$

$$\Delta G_{KH} = 2\pi\rho kT \sum_{s\alpha} \int_0^\infty [-2c_{s\alpha}(r) - h_{s\alpha}(r)(c_{s\alpha}(r) - \Theta(-h_{s\alpha}(r))h_{s\alpha}(r))] r^2 dr \quad (9)$$

where k is Boltzman constant, T is absolute temperature, ρ is bulk solvent number density, $\tilde{h}_{s\alpha}(r) = \sum_{s'\nu} \langle \tilde{\omega}_{ss'} * h_{s'\nu} * \tilde{\omega}_{\nu\alpha}^{\text{bulk}} \rangle$, $\tilde{\omega}_{ss'}$ and $\tilde{\omega}_{\nu\alpha}^{\text{bulk}}$ are the elements of matrices that are inverse to the matrices $\mathbf{W} = (\omega_{ss'})$ and $\mathbf{W}^{\text{bulk}} = (\omega_{\alpha\nu}^{\text{bulk}})$, which are built from the solute and solvent intramolecular correlation functions $\omega_{ss'}$ and $\omega_{\alpha\nu}^{\text{bulk}}$, respectively, $\Theta(x)$ is a Heaviside step function.

Note 3 *It should be mentioned, that although in RISM-MOL program All expressions are calculated with any choice of closure relation, some equations was derived for special closures. You may find closure- ΔG correspondance in the Table 1*

2.4 Excluded volume calculations and PWC expression

In the work [13] was proposed correction of the PW expression using the excluded volume term:

$$\Delta G_{PWC} = \Delta G_{PW} + a\rho V_{ex} \quad (10)$$

where a is empirical coefficient.

In the RISM-MOL program the excluded volume of the solute in infinitely dilute aqueous solution is calculated as a limiting case of the partial molar volume formula [14] when the solute density tends to zero:

$$V_{ex} = \frac{1}{\rho} + \frac{4\pi}{N_{solute}} \sum_s \int_0^\infty (h_{oo}^{bulk}(r) - h_{so}(r)) r^2 dr. \quad (11)$$

Here $h_{oo}^{bulk}(r)$ is the total oxygen-to-oxygen correlation function of bulk water, $h_{so}(r)$ is the total correlation function between the solute site s and the water oxygen.

2.5 Bridge correction and HNGB expression

The Hyper-Netted-Chain method with repulsive bridge correction (*HNGB*) was proposed by Kovalenko and Hirata in reference [6] has a modified form of this hydration free energy expression. The repulsive bridge functions $\{B_{s\alpha}^R(r)\}$ are defined for each pair of solute s and solvent α atoms by the expression:

$$\exp(-B_{s\alpha}^R(r)) = \prod_{\nu \neq \alpha} \left\langle \omega_{\alpha\nu}^{bulk} * \exp\left(-\beta \varepsilon_{s\nu} \left(\frac{\sigma_{s\nu}}{r}\right)^{12}\right) \right\rangle \quad (12)$$

where $\omega_{\alpha\nu}^{bulk}(r)$ are the solvent intramolecular correlation functions, and $\sigma_{s\nu}$ and $\varepsilon_{s\nu}$ are the site-site parameters of the pair-wise Lennard-Jones potential.

In the RISM-MOL program the repulsive bridge is implemented as an optional additional term to the potential:

$$u_{s\alpha}^{Full}(r) = u_{s\alpha}(r) + K_{br} \frac{1}{\beta} B_{s\alpha}^R(r) \quad (13)$$

where the constant K_{br} is either 0 (for HNC/PLHNC approximation) or 1 (for HNGB approximation). One should mention, that although the Bridge expansion was derived for HNC closure (4) only, in the RISM-MOL program there is a possibility to use it with the PLHNC closure (5) as well. To control, whether or not Bridge correction should be used, one may use the *user_UseBridgeCorrection* parameter in *UserParameters* file (see section 4.4 for details).

The free energy expression for the Repulsive Bridge is:

$$\Delta G_{HNGB} = \Delta G_{HNC} + 4\pi\rho kT \sum_{s\alpha} \int_0^\infty (h_{s\alpha}(r) + 1)(e^{-B_{s\alpha}^R(r)} - 1)r^2 dr \quad (14)$$

3 Algorithm

3.1 RISM equation reformulation

There are no known analytical methods to solve the set of RISM equations in general case. In the most cases RISM equations are solved numerically by iterative algorithms. The mostly

used method to solve the RISM equations is Fourier iterative scheme [15]. In this scheme the indirect correlation functions $\gamma_{s\alpha}(r) = h_{s\alpha}(r) - c_{s\alpha}(r)$ are introduced, and using the convolution theorem, the RISM equations (2) are rewritten in the Fourier space. To calculate the 3D Fourier transformations of the spherically symmetric functions and move back to the real space the forward and inverse Bessel-Fourier transformations $\mathcal{F}, \mathcal{F}^{-1}$ are used:

$$\mathcal{F}[f(r)] = \hat{f}(k) = \frac{4\pi}{k} \int_0^\infty r \sin(kr) f(r) dr \quad (15)$$

$$\mathcal{F}^{-1}[\hat{f}(k)] = f(r) = \frac{1}{2\pi^2 r} \int_0^\infty k \sin(kr) \hat{f}(k) dk \quad (16)$$

The RISM equations in the Fourier space are:

$$\hat{\gamma}_{s\alpha}(k) = \sum_{s'\nu} \hat{\omega}_{ss'}(k) \cdot \hat{c}_{s'\nu}(k) \cdot (\hat{\omega}_{\alpha\nu}^{\text{bulk}}(k) + \rho \hat{h}_{\alpha\nu}^{\text{bulk}}(k)) - \hat{c}_{s\alpha}(k) \quad (17)$$

where $\hat{\gamma}_{s\alpha}(k), \hat{c}_{s'\nu}(k), \hat{\omega}_{ss'}(k), \hat{\omega}_{\alpha\nu}^{\text{bulk}}(k), \hat{h}_{\alpha\nu}^{\text{bulk}}(k)$ are the Bessel-Fourier transformation of the functions $\gamma_{s\alpha}(r), c_{s'\nu}(r), \omega_{ss'}(r), \omega_{\alpha\nu}^{\text{bulk}}(r), h_{\alpha\nu}^{\text{bulk}}(r)$ respectively. The closure relation 3 is rewritten in terms of the indirect correlation functions:

$$c_{s\alpha}(r) = e^{-\beta u_{s\alpha}(r) + \gamma_{s\alpha}(r) + B_{s\alpha}(r)} - \gamma_{s\alpha}(r) - 1 \quad (18)$$

3.2 Short-range RISM formulation

Typically, functions $c_{s\alpha}(r)$ and $\gamma_{s\alpha}(r)$ decay slow with the distance. One need to consider them at large distances, which require large grids and computational time. Fortunately the asymptotic behaviour of the correlation functions is known. The direct correlation functions at the large distances can be well predicted by the Mean Spherical Approximation (MSA)

$$c_{s\alpha}(r) \approx -\beta u_{s\alpha}(r), \quad r > r_0 \quad (19)$$

where r_0 is a large enough constant. Because the Lennard-Jones potential decays fast at large distances we may approximate the asymptotic behavior of the direct correlation functions using only the Coulomb interaction potential. For the smooth splitting the functions into the long-range and short-range parts the Gauss error function $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz$ is usually used. We define the short-range and long-range potentials $u_{s\alpha}^S(r), u_{s\alpha}^L(r)$ as follows:

$$u_{s\alpha}^L(r) = \frac{q_s q_\alpha}{r} \text{erf}(tr) \quad u_{s\alpha}^S(r) = u_{s\alpha}(r) - u_{s\alpha}^L(r) \quad (20)$$

where $\frac{q_s q_\alpha}{r}$ is Coulomb potential, the parameter t controls the smoothness of transition between the short-range and long-range functions (typically $t = 1$). We define the long-range direct correlation functions $c_{s\alpha}^L(r)$ using the Mean Spherical Approximation:

$$c_{s\alpha}^L(r) = -\beta u_{s\alpha}^L(r) \quad (21)$$

Consistently, we define the short-range direct and indirect functions $c_{s\alpha}^S(r)$ and $\gamma_{s\alpha}^S(r)$ respectively as follows:

$$\begin{aligned} c_{s\alpha}^S(r) &= c_{s\alpha}(r) - c_{s\alpha}^L(r) = c_{s\alpha}(r) + \beta u_{s\alpha}^L(r) \\ \gamma_{s\alpha}^S(r) &= h_{s\alpha}(r) - c_{s\alpha}^S(r) = \gamma_{s\alpha}(r) - \beta u_{s\alpha}^L(r) \end{aligned} \quad (22)$$

The short-range functions are convenient for the numerical treatment of the problem. They decay rapidly with a distance and can be well approximated with the functions with the small support. Putting the short-range functions to the equation (18) we obtain the short-range closure relation:

$$c_{s\alpha}^S(r) = e^{-\beta u_{s\alpha}^S(r) - \gamma_{s\alpha}^S(r) + B_{s\alpha}(r)} - \gamma_{s\alpha}^S(r) - 1 \quad (23)$$

Following the reference [16] we write the analytical representation of the Bessel-Fourier transform of the long-range potential $u_{s\alpha}^L(r)$:

$$\hat{u}_{s\alpha}^L(k) = \frac{4\pi q_s q_\alpha}{k^2} e^{\frac{-k^2}{4t^2}} \quad (24)$$

One should mention, that this function decays fast with the k parameter, which makes it possible to use small grids in the Fourier space. Putting the short-range functions to the Fourier formulation of the RISM equation (17) we obtain the RISM equations for the short range functions:

$$\hat{\gamma}_{s\alpha}^S(k) = \sum_{s'\nu} \hat{\omega}_{ss'}(k) \cdot (\hat{c}_{s'\nu}^S(k) + \hat{u}_{s\alpha}^L(k)) \cdot (\hat{\omega}_{\alpha\nu}^{\text{bulk}}(k) + \rho \hat{h}_{\alpha\nu}^{\text{bulk}}(k)) - \hat{c}_{s\alpha}^S(k) \quad (25)$$

3.3 Fourier iterative algorithm

3.3.1 Iterations with the λ coupling parameter

To solve the RISM equations the RISM-MOL program uses the Fourier iterative method [15] speeded up by the multi-grid technique. [17] Fourier iteration process can be formally written in the form:

$$\gamma_{s\alpha}^{(n+1)} = \lambda \gamma_{s\alpha}^{(n)} + (1 - \lambda) \gamma'_{s\alpha}(r) \quad (26)$$

where $\gamma_{s\alpha}^{(n)}(r)$ and $\gamma_{s\alpha}^{(n+1)}(r)$ are indirect correlation function approximations on the iteration steps n and $n + 1$ respectively, $\lambda \in (0, 1]$ is the coupling parameter and $\gamma'_{s\alpha}(r)$ is an intermediate result of iterative procedure. In the next section will be discussed the procedure of calculating $\gamma'_{s\alpha}(r)$. And here we should put some remarks about the coupling parameter: For the physically-reasonable values of the interaction potential solution is stable. That means, that there exists a certain neighbourhood Ω of the fixed point such that for any initial guess $\Gamma^{(0)} \in \Omega$ the iterative process (4.4.3) with a small enough coupling parameter $\lambda \in (0, 1)$ converges to the solution of equations (2) Rigourously speaking, to warranty the convergence of the iterations on should beforehand know the neighbourhood Ω and use the coupling parameter $\lambda \rightarrow 0$. However, usually for the practical applications initial approximation $\gamma_{s\alpha}(r) = 0$ and $\lambda = \frac{1}{2}$ is enough to obtain the solution.

In the RISM-MOL program in the case of divergence of the algorithm, you may change the λ : it is controlled by the *user_LambdaCoupling* parameter in the *UserParameters* file (see section 4.4 for details).

3.3.2 Fourier iteration step

To obtain the $\gamma'_{s\alpha}(r)$ in the equation (4.4.3) from the functions $\gamma_{s\alpha}^{(n)}(r)$ the following actions are performed:

1. The direct correlation functions $c_{s\alpha}(r)$ are recalculated using the closure relation:

$$c_{s\alpha}(r) = e^{-\beta u_{s\alpha}(r) + \gamma_{s\alpha}(r) + B_{s\alpha}(r)} - \gamma_{s\alpha}(r) - 1 \quad (27)$$

where $u_{s\alpha}(r)$ is the site-site interaction potential and $B_{s\alpha}(r)$ is a site-site bridge function. In the current work the Partially-Linearized closure was used. [11, 12]

2. The Bessel-Fourier transformations $\hat{c}_{s\alpha}(k)$ of the functions $c_{s\alpha}(r)$ are found:

$$\hat{c}_{s\alpha}(k) = \frac{4\pi}{k} \int_0^\infty c_{s\alpha}(r) \cdot r \sin kr dr \quad (28)$$

3. The Fourier transformations $\hat{\gamma}_{s\alpha}(k)$ of the indirect correlation function are found using the relation (25)
4. The functions $\gamma'_{s\alpha}(r)$ are found as the inverse Bessel-Fourier transformations of the functions $\hat{\gamma}_{s\alpha}(k)$:

$$\gamma'_{s\alpha}(r) = \frac{1}{2\pi^2 r} \int_0^\infty \hat{\gamma}_{s\alpha}(k) \cdot k \sin kr dk \quad (29)$$

The iterative procedure might be presented in the schematic form:

$$\gamma_{s\alpha}^{(n)} \xrightarrow{\text{closure}} c_{s\alpha}(r) \xrightarrow{BFT} \hat{c}_{s\alpha}(k) \xrightarrow{eq.(??)} \hat{\gamma}_{s\alpha}(k) \xrightarrow{IBFT} \gamma'_{s\alpha}(r) \xrightarrow{eq.(??)} \gamma_{s\alpha}^{(n+1)} \quad (30)$$

where *BFT* and *IBFT* stands for *Bessel-Fourier Transform* and *Inverse Bessel-Fourier Transform* respectively.

3.3.3 Accuracy Theshold

Let $\Gamma^{(n)}$ be the set of the indirect correlation functions approximations on the step n :

$$\Gamma^{(n)} = \{ \gamma_{s\alpha}^{(n)}(r), s = 1 \dots N_{solute}, \alpha = 1 \dots N_{solvent} \} \quad (31)$$

One can formally define the iterative operator F , which implements the calculation scheme (30):

$$\Gamma^{(n+1)} = F\Gamma^{(n)} \quad (32)$$

The iterations are performed until the norm of difference between two subsequent solutions is less than the threshold ε . As the norm of the set of indirect correlation functions we may use the mean of the L_2 norms:

$$\|\Gamma^{(n+1)} - \Gamma^{(n)}\| \equiv \frac{1}{N_{solute}N_{solvent}} \sum_{s\alpha} \|\gamma_{s\alpha}^{(n+1)} - \gamma_{s\alpha}^{(n)}\|_{L_2} \leq \varepsilon \quad (33)$$

The smaller is ε , the more iteration steps will be performed and the better solution obtained. Typically default value $\varepsilon = 10^{-4}$. If you need more precise calculations, you may change ε by changing the *user_AccuracyThreshold* parameter in the *UserParameters.m* file (see section 4.4 for details)

3.4 Multi-Grid

To improve the efficiency of the iteration algorithm in the RISM-MOL program the multi-grid technique is used. The multi-grid technique operates with the solution approximations discretized on the grid with different number of points. If one grid has less points than another, then the first one is called "coarse grid" and the second one is called "fine grid". The grid with the smallest number of points is called "the coarsest grid". The idea of the multi-grid technique is to find the low-frequency components of the solution on the coarse grids and use the fine grids to find the high-frequency perturbations of the coarse-grid solution. Also the multi-grid accounts the fact, that due to discretization of the iterative operators, iterative processes on different grids converge to different solutions.

Let Γ_{N-k} be the approximate solution discretized on the grid with 2^{N-k} points and step $2^k \Delta R$, $0 \leq k \leq L$. Let for the discretized solutions two operators be defined:

- Interpolation operator P interpolates the solution to the grid with the twice larger number of points and twice smaller grid step:

$$\Gamma'_N = P\Gamma_{N-1} \quad (34)$$

- Restriction operator R removes each second grid point from the solution:

$$\Gamma_{N-1} = R\Gamma_N \quad (35)$$

Let F_{N-k} be the describe iterative operator on the grid with 2^{N-k} points, $0 \leq k \leq L$. Due to the discretization errors, the restricted exact solution on a fine grid is not the solution on the coarse grid. Let Γ_{N-k}^{exact} be the exact solution on the grid with 2^{N-k} points:

$$\Gamma_{N-k}^{exact} = F_{N-k}\Gamma_{N-k}^{exact} \quad (36)$$

Then

$$R\Gamma_{N-k}^{exact} \neq F_{N-k-1}R\Gamma_{N-k}^{exact} \quad (37)$$

The difference between the exact solution due to the operator discretization can be approximated by the difference of approximate solutions:

$$\Delta\Gamma_{N-k-1}^{(n)} = RF_{N-k}\Gamma_{N-k}^{(n)} - F_{N-k-1}R\Gamma_{N-k}^{(n)} \quad (38)$$

To correct the error due to the operator discretization this difference should be added to the iterative operator on the coarse grid:

$$\Gamma_{N-k-1}^{(n+1)} = F_{N-k-1}\Gamma_{N-k-1}^{(n)} + \Delta\Gamma_{N-k-1}^{(n)} \equiv \left(F_{N-k} + \Delta\Gamma_{N-k}^{(n)}\right)\Gamma_{N-k}^{(n)} \quad (39)$$

Calculation of the multi-grid iterative operator $\mathcal{M}_{N-k}(\Gamma_{N-k}, \Delta\Gamma_{N-k})$ on the grid with 2^{N-k} points with the grid correction $\Delta\Gamma_{N-k}$ includes the following steps:

1. Perform small number ν of smoothing iteration on the fine grid (typically $\nu_1 \leq 5$):

$$\Gamma_{N-k}^{(\nu_1)} = (F_{N-k} + \Delta\Gamma_{N-k})^{\nu_1}\Gamma_{N-k} \quad (40)$$

2. Calculate the correction for the grid with 2^{N-k-1} grid points

$$\Delta\Gamma_{N-k-1} = RF_{N-k}\Gamma_{N-k}^{(\nu_1)} - F_{N-k-1}R\Gamma_{N-k}^{(\nu_1)} \quad (41)$$

3. **if** $N - k - 1 = N - L$ **then** perform n Fourier iterations on the coarsest grid:

$$\Gamma_{N-L}^{(1)} = (F_{N-L} + \Delta\Gamma_{N-L})^n R\Gamma_{N-k} \quad (42)$$

else perform μ multi-grid iterations on the grid with 2^{N-k-1} points (typically $\mu \leq 2$)

$$\Gamma_{N-k-1}^{(1)} = \mathcal{M}_{N-k-1}^\mu(R\Gamma_{N-k}^{\nu_1}, \Delta\Gamma_{N-k-1}) \quad (43)$$

4. Use the solution on the grid with 2^{N-k-1} points to refine the solution on the grid with 2^{N-k} points:

$$\Gamma_{N-k}^{(\nu_1+1)} = \Gamma_{N-k}^{(\nu_1)} + P(\Gamma_{N-k-1}^{(1)} - R\Gamma_{N-k}^{\nu_1}) \quad (44)$$

5. Perform ν_2 smoothing fine-grid iterations (typically $\nu_2 \leq 5$)

$$\Gamma_{N-k}^{(\nu_1+1+\nu_2)} = (F_{N-k} + \Delta\Gamma_{N-k})^{\nu_2} \Gamma_{N-k}^{\nu_1+1} \quad (45)$$

The solution approximation $\Gamma_{N-k}^{\nu_1+1+\nu_2}$ is the result of the multi-grid iteration:

$$\Gamma_{N-k}^{(\nu_1+1+\nu_2)} = \mathcal{M}_{N-k}(\Gamma_{N-k}, \Delta\Gamma_{N-k}) \quad (46)$$

It was shown previously that the multi-grid solver speeds up the Fourier iterations for the atomic Ornstein-Zernike equation up to several dozen times. [18] The RISM-MOL program gives the speed up up to 30 times for the molecular RISM equations.

4 Algorithm Parameters

4.1 Site-site potential description

To calculate the closure relation (3) one need to know the site-site interaction potential $u_{s\alpha}(r)$. In the RISM-MOL the interaction potential is calculated as sum of three terms: Lennard-Jones (LJ) potential, Coulomb potential and optional repulsive bridge potential (see section 2.5)

$$u_{s\alpha}(r) = u_{s\alpha}^{LJ}(r) + u_{s\alpha}^C(r) + K_{br} \frac{1}{\beta} B_{s\alpha}^R(r) \quad (47)$$

where $u_{s\alpha}^C(r)$ is the Coulomb potential

$$u_{s\alpha}^C(r) = \frac{q_s q_\alpha}{r} \quad (48)$$

$u_{s\alpha}^{LJ}(r)$ is a Lennard-Jones potential

$$u_{s\alpha}^{LJ}(r) = 4\epsilon_{s\alpha} \left(\left(\frac{\sigma_{s\alpha}}{r} \right)^{12} - \left(\frac{\sigma_{s\alpha}}{r} \right)^6 \right) \quad (49)$$

$B_{s\alpha}^R(r)$ is a repulsive bridge (12), constant $K_{br} \in \{0, 1\}$ determines whether or not bridge correction should be used.

4.1.1 Lennard Jones Mixing rules

The pair Lennard-Jones parameters $\sigma_{s\alpha}$, $\epsilon_{s\alpha}$ are calculated via the combining rules. In RISM-MOL program two mixing rules are implemented:

1. Lorentz-Berthelot rules (default):

$$\sigma_{s\alpha} = \frac{\sigma_s + \sigma_\alpha}{2} \quad \epsilon_{s\alpha} = \sqrt{\epsilon_s \epsilon_\alpha} \quad (50)$$

2. OPLSAA geometric rules

$$\sigma_{s\alpha} = \sqrt{\sigma_s \sigma_\alpha} \quad \epsilon_{s\alpha} = \sqrt{\epsilon_s \epsilon_\alpha} \quad (51)$$

To choose, which rules will be used in the program one may use the parameters *user_MixingRules* in the file *UserParameters.m* (see section 4.4 for details)

4.2 Parameters of solvent

In the current version of the program the following parameters of the solvent are used:

- temperature $T = 300K$
- number density $n = 0.0337$ molecules/ \AA^3 (33.7 molecules/ nm^3)

4.3 Parameters of the solute molecule

In the RISM method solute molecule is considered as a rigid structure of atoms, which stay at the constant distances each from other. Among the interactions are considered only coulomb and Van der Waals interactions. To describe Van der Waals interactions Lennard Jones potentials are used. Coulomb interactions depends on charges of atoms. In such approximations solute molecule of N_{solute} atoms might be described by such parameters:

1. N_{solute} triples of atom coordinates

$$(x_i, y_i, z_i) \quad i = 1..N_{solute}$$

2. N_{solute} pairs of Lennard-Jones parameters of atoms

$$(\sigma_i, \epsilon_i) \quad i = 1..N_{solute}$$

3. N_{solute} partial charges of atoms

$$q_i \quad i = 1..N_{solute}$$

4.3.1 Format of input solute structure file

Input data for the solute is stored in a separate file. Structure file is an regular text file. There should be six space or tab separated columns of numbers in it. Number of rows is equal to the number of atoms in the solute molecule. Each row corresponds to the one atom.

Data columns are: **X Y Z Sigma Epsilon Charge**

- **X,Y,Z** are coordinates of the atom
- **Sigma** and **Epsilon** are Lenard-Johnes potential parameters for the atom
- **Charge** is the charge of the atom

X Y Z and **Sigma** should be given in the input distance units. **Epsilon** should be given in the input Energy units. **Charge** should be given in the atomic charge units (1 a.u. charge = charge of proton).

By default input and output distance units are angstroms and energy units kcal/mol. However there is possibility to write an input file in a different units and specify these units as parameters of the function *startRISM* (see section ?? for detailed information).

4.3.2 Example of the input structure file

Here is the example of solute structure for methane molecule with OPLSAA force-filed parameters:

Example 1 Example of the input solute structure file

-0.4458	-0.0120	0	3.5	0.066	-0.24
-0.0823	-1.0402	0	2.5	0.030	0.06
-0.0823	0.5020	0.8904	2.5	0.030	0.06
-0.0823	0.5020	-0.8904	2.5	0.030	0.06
-1.5363	-0.0120	0	2.5	0.030	0.06

The first line corresponds to the Carbon atom, lines 2-5 describe hydrogen atoms. Coordinates (columns 1-3) and σ -parameters(4-th column) are given in Angstroems. ϵ parameters (5-th colums) are given in kcal/mol. Charge (the last column) are given in proton charges.

4.4 UserParameters file

In the RISM-MOL it is possible to change the water parameters, use different closure relations, Lenneard Jones mixing rules and numerical scheme parameters. To adjust these parameters one may edit the *UserParameters.m* file, which is located in the folder, where RISM-MOL is installed. If you have successfully setup paths to the RISM-MOL program, you can edit this file from the Matlab by the command:

edit UserParameters The file contain the following sections:

- Water parameters
- Model parameters
- Numerical Scheme parameters

The units used in file are:

- **Distance units:** Å
- **Energy units:** kcal/mol
- **Charge units:** proton charge

4.4.1 Water parameters

By default in the RISM-MOL modified SPC/E water model (MSPC/E) is used. [19] However, there is a possibility to change the model. Rigid tree-point water molecule model is defined by oxygen and hydrogen LJ parameters and partial charges, and by distances between the atoms.

Charges of water oxygen and hydrogen To change the charges of water oxygen and hydrogen you may change *user_HydrogenCharge* and *user_OxygenCharge* parameters in the *UserParameters.m* file.

Lennard Jones parameters of water atoms To change σ and ϵ LJ parameters of water oxygen and hydrogen you may change *user_OxygenSigma*, *user_OxygenEpsilon*, *user_HydrogenSigma*, *user_HydrogenEpsilon* parameters in the *UserParameters.m* file.

Note 4 • *Water Hydrogen σ parameter strongly affects the value of Hydration Free Energy.*

- *Small values of Hydrogen σ can lead to divergence of the algorithm, when HNC closure is used*

Distances between the atoms Water molecule is supposed to be symmetric. The structure of molecule can be defined by two distances: Length of the O-H bond and distance between the Hydrogen atoms. These distances can be changed by changing the *user_DistanceOxygenHydrogen* and *user_DistanceHydrogenHydrogen* parameters in the *UserParameters.m* file.

4.4.2 Model parameters

In the RISM-MOL user may change the parameters of the RISM model used for calculation. Such parameters are :

- Choice of closure relation (see section 2.2 for details)
- Choice of the Lennard Jones Mixing rules (see section 4.1.1 for details)
- Whether or not Repulsive Bridge is used (see section 2.5 for details)

Closure Closure Relation is controlled by the *user_Closure parameter* There are two possibilities for closure relation:

- *user_Closure* = 'HNC' - Hyper Netted Chain Closure (4)
- *user_Closure* = 'PLHNC' - Partially Linearized Closure (5) (default)

Lennard Jones Potential Mixing rules In the RISM-MOL two different mixing rules are implemented to obtain pair potential parameters from atom parameters:

- Geometric OPLSAA rules (51)

To use these rules set: `user_MixingRules = 'OPLSAA';`

- Lorentz-Berthelot rules (50) (default)

To use these rules set: `user_MixingRules = 'LorentzBerthelot';`

Repulsive Bridge Correction In the RISM-MOL there is a possibility to turn on or turn off the repulsive bridge correction (12) by affecting parameters K_{br} in the potential (47). To indicate, whether or not bridge correction should be used one may change the `user_UseRepulsiveBridge` parameter in the *UserParameters* file. Possible values:

- `user_UseRepulsiveBridge = 'no'` (default). Turn off Bridge correction (set $K_{br} = 0$)
- `user_UseRepulsiveBridge = 'yes'`. Turn on Bridge correction (set $K_{br} = 1$)

4.4.3 Numerical Scheme parameters

In the RISM-MOL there is a possibility to choose the parameters of the iterative solver. One may change the *Coupling parameter* λ (see section 3.3.1), and choose the accuracy threshold ε for iterations (see section 3.3.3).

Iterations Coupling parameter λ The λ parameter determines the size of the iterative step in the iterative process. Valid values for lambda: $0 < \lambda \leq 1$. The larger is λ , the faster iterations converge. However, the larger is λ , the larger is the possibility, that iterations will diverge. So, one should choose the λ coupling parameter carefully. Typically, for the default PLHNC closure (5) iterations almost always converge with the default value $\lambda = 0.5$. However, if you use HNC closure (4), or repulsive bridge (12), iterations may diverge for highly polarizable molecules. In that case it is necessary to decrease λ parameter.

To change the λ coupling parameter use `user_LambdaCoupling` parameter in the *UserParameters.m* file.

Accuracy Threshold to stop the iterations Accuracy threshold ε determines the condition, when the iterations will stop. Iterations stops, when the difference between two sequential approximations (33) is less than ε . The smaller accuracy threshold is, the better accuracy of solutions you obtain. Typically, the standard value $\varepsilon = 10^{-4}$ is enough. It corresponds to 0.02 kcal/mol error in the free energy calculations for simple organic molecules. However, if you need better accuracy, you may choose smaller threshold.

To change the accuracy threshold ε use `user_AccuracyThreshold` parameter in the *UserParameters.m* file.

5 Program interface: startRISM function

The interface with a program is made by the only matlab function - startRISM. If you have setup the paths to the program properly, you can run this function from the Matlab command window

Function takes as input the coordinates and force-field parameters of the solute molecule from the Input solute file (see section 4.3.1) and calculates hydration free energies by different expressions, (see section 2.3), Excluded volume (11), site-site Radial Distribution Functions $g_{s\alpha}(r)$, solutions of the RISM equations (25) $\gamma_{s\alpha}^S(r)$ and $c_{s\alpha}^S(r)$ and long-range potential $u_{s\alpha}^L(r)$ defined by eq. (20).

To run the **startRISM** function use the following command in Matlab Command Window:

```
[Mu,g,grid,gamma, c,potLong] = startRISM (fname, input_units, output_units,xparam)
```

Here Mu , g , $grid$, $gamma$, c , $potLong$ are output parameters $fname$, $input_units$, $output_units$, $xparam$ are input parameters

For Example you may use such command:

```
[Mu, g,grid,gamma, c,potLong] = startRISM ('methane.txt', ...
    struct('Distance','Angstr','Energy','kcal/mol'), ...
    struct('Distance','Angstr','Energy','kcal/mol'), ...
    'user.LambdaCoupling=0.5;')
```

You may also omit last some input argument, to use the default parameters, and some output arguments, if you don't need them:

```
[Mu,g,grid] = startRISM ('methane.txt')
```

The meaning of the parameters, and references to their detailed description:

Input parameters

- $fname$ - path to the input file with the solute structure. see subsection 5.1.1
- $input_units$ - units, used for the coordinates in the structure file and for σ LJ parameters (by default - Å) see subsection 5.1.2
- $output_units$ - units, used for calculated Hydration Free Energies and distance vector $grid$. see subsection 5.1.2
- $xparam$ - string, containing extra parameters. Any parameters from *UserParameters* file can be used here. see subsection 5.1.5.

Output parameters

- Mu - structure, which contains calculated hydration free energies and excluded volume. see subsection 5.2.1.
- g - site-site radial distribution functions $\{g_{s\alpha}(r)\}$. see subsection 5.2.3
- $grid$ - grid vector. see subsection 5.2.3
- $gamma, c$ - short-range solutions $\gamma_{s\alpha}^S(r)$, $c_{s\alpha}^S(r)$ of the RISM equations (25). see subsection 5.2.5
- $potLong$ - long range part of potential $\beta u_{s\alpha}^L(r)$ see subsection (5.2.5)

5.1 Input Parameters of startRISM function

5.1.1 fname

Path to the input structure file. This file describes the structure of the solute molecule. Format of this file is described in the section 4.3.1

5.1.2 input_units,output_units

input_units define the units in which the input file is written. **output_units** define the units in which the chemical potentials and grids are produced.

Both parameters have the same format. They are the structures, with two fields:

- *Distance* – the name of the distance units
- *Energy* – the name of the energy units
- **input_units.Distance** are the units, used for the **X,Y,Z** and **Sigma** columns of the solute input file
- **input_units.Energy** are the units, used for the **Epsilon** column of the solute input file
- **output_units.Distance** are the units, used for the **grid** output parameter.
- **Mu.VUA** (excluded volume) is measured in **output_units.Distance**³
- **output_units.Energy** are the units, used for the hydration free energy

5.1.3 Supported units

Distance Units

- 'm' – meters
- 'Angstr' – Angstroms ($1 \text{ \AA} = 10^{-10} \text{ m}$)
- 'Bohr' – Atomic distance units (Bohr) ($1 \text{ Bohr} = 5.291772108 \cdot 10^{-11} \text{ m}$)

Energy Units

- 'J' – joules
- 'Hartree' – Hartree , atomic energy units ($1 \text{ Hartree} = 4.35974417 \cdot 10^{-18} \text{ J}$)
- 'kcal/mol' – kilocalories per mole. ($1 \text{ kcal/mol} = 6.9477 \cdot 10^{-21} \text{ J}$)
- 'KBT' – Boltzman constant multiplied by the temperature. The value is given for T=298K (25 C), $1 \text{ KBT} = 4.11 \cdot 10^{-21} \text{ J}$
- 'kJ/mol' – kilojoules per mole, ($1 \text{ kJ/mol} = 0.239 \text{ kcal/mol} = 1.6605003 \cdot 10^{-21} \text{ J}$)
- 'eV' – electron Voltes ($1 \text{ eV} = 1.60217653 \cdot 10^{-19} \text{ J}$)

Parameters **input_uniots**, **output_units** are optional and can be omitted.

In that case the default values will be used: Å for distance and kcal/mol for energy

5.1.4 How to define input/output units

To define the input/output parameters type in matlab command window:

```
input_parameters = struct('Distance',distance unit names, 'Energy',energy unit names)
```

Example 2 Define input/output parameters for startRISM

```
inputParameters = struct('Distance','Angstr','Energy','kcal/mol')
outputParameters = struct('Distance','Angstr','Energy','kcal/mol')
Mu=startRISM('methane.txt',inputParameters,outputParameters)
```

5.1.5 Extra parameters: xparam

Sometimes you need to change some parameters of the algorithm (e.g closure relation, or accuracy). As was said before, it can be done by editing the *UserParameters.m* file. However, if you need to make a lot of calculations with *different* parameters it might be not convenient to change *UserParameters.m* many times. Moreover, it may lead to mistakes. That's why in RISM-MOL was implemented the possibility to change parameters *on fly*, just for the current run of the program. To do it you should use *xparam* parameter of the **startRISM** function.

xparam is the string containing the values of the algorithm parameters, separated by semicolon (;). Parameters are given in the format

Parameter name = Parameter value

For the list of available parameters you may refer the section 4.4 (parameters are the same, which are used in the *UserParameters* file).

Note 5 If the parameter value is string, you should use two **single** quotes instead each quote. For example: to change the closure to HNC use the following value of *xparam*:

```
xparam='user_Closure="HNC";'
```

5.1.6 Example of using extra parameters

For example, you need to run the RISM program for the methane molecule, HNC closure, and to prevent the divergence you would like to use coupling parameter $\lambda = 0.05$. You may do it by typing following commands in the matlab command window (see example 3):

Example 3 Example of using extra parameters

```
inputParameters = struct('Distance','Angstr','Energy','kcal/mol')
outputParameters = struct('Distance','Angstr','Energy','kcal/mol')
xparam='user_Closure="HNC";user_LambdaCoupling=0.05;'
Mu=startRISM('methane.txt',inputParameters,outputParameters,xparam)
```

5.2 Output parameters of startRISM function

5.2.1 Mu: Hydration Free Energy and excluded volume

This parameter contains values of the chemical potentials, calculated by different expressions, and the value of the excluded volume. You may refer section 2.3 for the description of different expressions. Output parameter *Mu* is a structure with such fields:

- Mu.HNC – Hydration Free Energy, calculated using Hyper-Netted-Chain assumption (6).
- Mu.G – Hydration Free Energy, calculated using Gaussian Fluctuations expression (7)
- Mu.HNCB – Hydration Free Energy, calculated using the HNCB expression. (14)
- Mu.PW – Hydration Free Energy, calculated using Partial Wave expression (8)
- Mu.PWC – Hydration Free Energy, calculated using Partial Wave expression with excluded volume correction (10)
- Mu.VUA – excluded volume (11)
- Mu.KH – Hydration Free Energy, calculated using Kovalenko-Hirata expression for PLHNC closure (9)

Hydration Free Energies (**Mu.HNC**,**Mu.G**,**Mu.HNCB**,**Mu.PW**,**Mu.PWC**,**Mu.KH**) are in the **output_units.Energy** units (by default: kcal/mol),

Excluded volume (**Mu.VUA**) is in **output_units.Distance**³ units (by default Å³)

5.2.2 Example of Hydration Free Energies calculation

You may calculate Hydration Free Energies of methane by following matlab script:

```
Mu=startRISM('methane.txt');
```

To see values of the Hydration Free Energies use *disp* command:

```
disp(Mu)
```

To use Hydration Free Energy, calculated by some particular expression use *Mu.expression namej*. For example:

```
disp(Mu.HNC)
```

will show Hydration Free Energy, calculated by the HNC expression (6)

5.2.3 Radial distribution functions: *g* and *grid* parameters

Output parameter *g* This parameter contains site-site Radial Distribution Functions, calculated by RISM algorithm. The radial distribution functions are stored in the matrix. Each column of the matrix contain one distribution function.

Odd columns (1,3,5,...) contain the distribution functions between the solute sites and water oxygen. Column with number $2k + 1$ contain the distribution functions $g_{kO}(r)$ between

the k -th solute atom and water oxygen (solute atoms are numbered in the same order, as they give in the *Solute Input File*. See esection 4.3.1 for details).

Even columns (2,4,6,...) contain the distribution functions between the solute sites and water hydrogen. Column with number $2k + 2$ contain the distribution functions $g_{kH}(r)$ between the k -th solute atom and water hydrogen (solute atoms are numbered in the same order, as they give in the *Solute Input File*. See esection 4.3.1 for details).

You may access particular RDF function by using the matlab colon operator ':' For example $g(:,3)$ will return the RDF function between the 2-ns solute site and water oxygen (because $3 = 2 \cdot 1 + 1$).

Output parameter *grid* It's the column-array, containing the samples of distance from the solute atom. That samples corresponds to the samples in the produced RDF functions (columns of g array). $g(i,k)$ is the value of the k -th RDF function at the distance $grid(i)$

5.2.4 Example of calculating RDF functions

Let for example we would like to calculate site-site RDF functions between the atoms of methane molecule and water. We may calculate them by typing such line in the Matlab Command Window:

```
[Mu,g,grid]=startRISM('methane.txt')
```

To access particular RDF function we use the matlab colon operator ':'. The Input file for the methane molecule you may find in the example 1. As you remember, in that file the first line corresponds to the Carbon atom, other lines correspond to hydrogen. Then, to plot the Carbon-Water oxygen and Carbon-water hydrogen RDF functions we may use the following commands:

```
plot(grid,g(:,1),grid,g(:,2))
```

To make them more recognizable we use additional visualization commands:

```
legend('C-O','C-H');
xlim([0 10]);
xlabel('Angstroms');
ylabel('g(r)');
```

5.2.5 Solutions of the RISM equations: *gamma,c* and *potLong* parameters

To obtain the solutions of the RISM equations (25) you may use the *gamma,c* and *potLong* parameters.

gamma is the matrix, containing the short-range site-site indirect correlation functions $\gamma_{s\alpha}(r)$ c is the matrix, containing the short-range site-site direct functions $c_{s\alpha}(r)$. (see formulae (22)) *potLong* is the matrix, containing the long-range potential $\beta u_{s\alpha}^L(r)$ (see (20)) The format of the matrices as the same, as for the site-site RDF functions (subsection 5.2.3)

To obtain full direct and indirect correlation functions $\gamma_{s\alpha}(r)$ and $c_{s\alpha}(r)$ one may use equations:

$$\begin{aligned}\gamma_{s\alpha}(r) &= \gamma_{s\alpha}^S(r) + \beta u_{s\alpha}^L(r) \\ c_{s\alpha}(r) &= c_{s\alpha}^S(r) - \beta u_{s\alpha}^L(r)\end{aligned}\tag{52}$$

5.2.6 Example of working with gamma, c , and potLong parameters

To obtain the solutions of the RISM equations for methane molecule you may use the following script:

```
[Mu,g,grid,gammaS,cS,potLong]=startRISM('methane.txt');
```

To plot short-range functions you may use the script

```
figure,plot(grid,gammaS), xlim([0 10]), title('Gamma functions')
figure, plot(grid,cS), xlim([0 10]), title('c functions')
```

To plot the long-range potential you may use:

```
figure,plot(grid,potLong), xlim([0 10]), title('\beta u_{Long}')
```

To find full functions you may use the script:

```
c_full = cS - potLong; gamma_full = gammaS + potLong;
```

To plot the full functions you may use the script:

```
figure,plot(grid,gamma_full), xlim([0 10]), title('Gamma functions')
figure, plot(grid,c_full), xlim([0 10]), title('c functions')
```

References

- [1] F. Hirata, editor. *Molecular theory of solvation*. Kluwer Academic Publishers, Dordrecht, Netherlands, 2003.
- [2] J.-P. Hansen and I. R. McDonald. *Theory of Simple Liquids, 4th ed.* 2000.
- [3] M. Kinoshita, Y. Okamoto, and F. Hirata. Calculation of solvation free energy using RISM theory for peptide in salt solution. *Journal of Computational Chemistry*, 19(15):1724–1735, November 1998.
- [4] M. Kinoshita, Y. Okamoto, and F. Hirata. Calculation of hydration free energy for a solute with many atomic sites using the rism theory: A robust and efficient algorithm. *Journal of Computational Chemistry*, 18(10):1320–1326, July 1997.
- [5] S. J. Singer and D. Chandler. Free-energy functions in the extended rism approximation. *Molecular Physics*, 55(3):621–625, 1985.
- [6] A. Kovalenko and F. Hirata. Hydration free energy of hydrophobic solutes studied by a reference interaction site model with a repulsive bridge correction and a thermodynamic perturbation method. *Journal of Chemical Physics*, 113:2793–2805, 2000.
- [7] D. Chandler, Y. Singh, and D. M. Richardson. Excess electrons in simple fluids .1. general equilibrium-theory for classical hard-sphere solvents. *Journal of Chemical Physics*, 81(4):1975–1982, 1984.

- [8] S. Ten-no. Free energy of solvation for the reference interaction site model: Critical comparison of expressions. *Journal of Chemical Physics*, 115(8):3724–3731, August 2001.
- [9] K. Sato, H. Chuman, and S. Ten-no. Comparative study on solvation free energy expressions in reference interaction site model integral equation theory. *Journal of Physical Chemistry B*, 109(36):17290–17295, September 2005.
- [10] S. Ten-no and S. Iwata. On the connection between the reference interaction site model integral equation theory and the partial wave expansion of the molecular Ornstein-Zernike equation. *Journal of Chemical Physics*, 111(11):4865–4868, 1999.
- [11] A. Kovalenko and F. Hirata. Potential of mean force between two molecular ions in a polar molecular solvent: A study by the three-dimensional reference interaction site model. *Journal of Physical Chemistry B*, 103:7942–7957, 1999.
- [12] A. Kovalenko and F. Hirata. Self-consistent description of a metal-water interface by the kohn-sham density functional theory and the three-dimensional reference interaction site model. *Journal of Chemical Physics*, 110:10095–10112, 1999.
- [13] G.N. Chuev, M.V. Fedorov, and J. Crain. Improved estimates for hydration free energy obtained by the reference interaction site model. *Chemical Physics Letters*, 448(4-6):198–202, 2007.
- [14] J. G. Kirkwood and F. P. Buff. The statistical mechanical theory of solutions .1. *Journal of Chemical Physics*, 19(6):774–777, 1951.
- [15] P. A. Monson and G. P. Morriss. Recent progress in the statistical-mechanics of interaction site fluids. *Advances in Chemical Physics*, 77:451–550, 1990.
- [16] K. C. Ng. Hypernetted chain solutions for the classical one-component plasma up to $\gamma = 7000$. *Journal of Chemical Physics*, 61(7):2680–2689, 1974.
- [17] W. Hackbusch. *Multi-grid methods and Applications*. Springer-Verlag, Berlin, 1985.
- [18] M. V. Fedorov and W. Hackbusch. A multigrid solver for the integral equations of the theory of liquids. Preprint 88, Max-Planck-Institut fuer Mathematik in den Naturwissenschaften, 2008.
- [19] L. Lue and D. Blankschtein. Liquid-state theory of hydrocarbon water-systemsapplication to methane, ethane, and propane. *Journal of Physical Chemistry*, 96(21):8582–8594, October 1992.